



Strategic COVID-19 vaccine distribution can simultaneously elevate social utility and equity

Lin Chen ^{1,2,10}, Fengli Xu ^{3,4,10} ✉, Zhenyu Han ², Kun Tang⁵, Pan Hui ^{1,6,7,8} ✉, James Evans ^{3,9} ✉ and Yong Li ² ✉

Balancing social utility and equity in distributing limited vaccines is a critical policy concern for protecting against the prolonged COVID-19 pandemic and future health emergencies. What is the nature of the trade-off between maximizing collective welfare and minimizing disparities between more and less privileged communities? To evaluate vaccination strategies, we propose an epidemic model that explicitly accounts for both demographic and mobility differences among communities and their associations with heterogeneous COVID-19 risks, then calibrate it with large-scale data. Using this model, we find that social utility and equity can be simultaneously improved when vaccine access is prioritized for the most disadvantaged communities, which holds even when such communities manifest considerable vaccine reluctance. Nevertheless, equity among distinct demographic features may conflict; for example, low-income neighbourhoods might have fewer elder citizens. We design two behaviour-and-demography-aware indices, community risk and societal risk, which capture the risks communities face and those they impose on society from not being vaccinated, to inform the design of comprehensive vaccine distribution strategies. Our study provides a framework for uniting utility and equity-based considerations in vaccine distribution and sheds light on how to balance multiple ethical values in complex settings for epidemic control.

Sweeping the globe since its outbreak, the COVID-19 pandemic continues to impose harm and loss on human societies everywhere¹. This has revealed and intensified disparities in the health conditions of different communities because of systemic inequities in COVID-19 exposure and access to health systems^{2–4}. In the fight against the pandemic, vaccines are regarded as the most critical medical resources, but they still face prominent shortages in many countries and communities^{5,6}. As a result, vaccine prioritization has become a critical policy task in every public health system^{7–9}, with well-considered strategies to balance multiple ethical values. Facing the pandemic as a shock to the whole society, we argue that all members should have an equal right and opportunity to attain the best protection from the pandemic. In this light, our aim is to achieve a desirable balance between social utility and equity, where social utility is defined as the prevention of mortality in the entire population, and equity is defined as the mitigation of mortality disparities in disadvantaged demographic groups^{10–13}. These two goals represent the most visible metrics considered by health authorities and organizations worldwide^{9,14–16}, undergirded by the contrasting ethical values of utilitarianism and egalitarianism. As J. Bentham put it, “the well-being of a portion of individuals” can sometimes be sacrificed to achieve “the greatest happiness of the greatest number”¹⁷. Previous research has identified trade-offs between social utility and equity in the distribution of health-care resources ranging from disease screening to treatment^{18–22}. In COVID-19 vaccine distribution, the most recent studies have focused on the trade-off between minimizing the years of life lost and minimizing the number of lives lost^{23–27}, both of which reflect social-utility-oriented

values and neglect disparities across the population. In light of this gap, we aim to reveal the relationship between social utility and equity in COVID-19 vaccine distribution, with critical implications for designing vaccine prioritization.

Examining the social utility and equity of vaccine distribution strategies requires an epidemiological model that can capture the uneven risks faced by different communities²⁶—for example, older persons and those with greater mobility have higher COVID-19 risk^{3,28–30}. However, standard epidemiological models (for example, susceptible–infectious–recovered (SIR)^{31,32} and susceptible–exposed–infectious–recovered (SEIR)^{33,34}) are built on the assumption of homogeneous population mixing, which prohibits them from capturing heterogeneity in the spread of coronavirus. Some recent work^{35–37} has aimed to augment standard epidemiological models with empirical mobility data, but these neglect inherent vulnerability differences embedded in demographic profiles²⁸. Here we propose an epidemiological model that simultaneously captures heterogeneity in the mobility patterns and demographic profiles of different communities. Calibrated with large-scale mobility and census data covering more than 75 million residents in the United States, our model automatically tunes the dynamics of coronavirus spread within each community on the basis of its demographic profile and traces dispersion among communities with time-varying empirical mobility flows. Our model can accurately predict the number of daily deaths and reconstruct its uneven distribution among communities, enabling the evaluation of equity among communities defined by different demographics—that is, older adult ratio, average household income, essential worker ratio and racial-ethnic minority ratio,

¹Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong SAR, P. R. China. ²Beijing National Research Center for Information Science and Technology (BNRist), Department of Electronic Engineering, Tsinghua University, Beijing, P. R. China.

³Knowledge Lab & Department of Sociology, University of Chicago, Chicago, IL, USA. ⁴Mansueto Institute for Urban Innovation, University of Chicago, Chicago, IL, USA. ⁵Vanke School of Public Health, Tsinghua University, Beijing, P. R. China. ⁶Computational Media and Arts Thrust, Hong Kong University of Science and Technology (Guangzhou), Guangzhou, P. R. China. ⁷Division of Emerging Interdisciplinary Area, Hong Kong University of Science and Technology, Hong Kong SAR, P. R. China. ⁸Department of Computer Science, University of Helsinki, Helsinki, Finland. ⁹Santa Fe Institute, Santa Fe, NM, USA. ¹⁰These authors contributed equally: Lin Chen, Fengli Xu. ✉e-mail: fenglixu@uchicago.edu; panhui@ust.hk; jevens@uchicago.edu; liyong07@tsinghua.edu.cn

which have garnered considerable attention throughout the pandemic^{4,28–30,35,38,39}. Accordingly, we examine four vaccine distribution strategies that prioritize communities on the basis of vulnerabilities defined by these four demographic features. We find that social utility and equity can be simultaneously improved when prioritizing the most disadvantaged communities in each demographic dimension for vaccine access. Such a result holds even when low-income communities show considerable vaccine hesitancy, which contrasts with the conventional view of inevitable trade-offs^{19–22}. Nevertheless, elevating the equity across one demographic dimension can degrade it across others, suggesting that demographic features alone are insufficient to guide vaccine distribution. To overcome this, we propose two demography-and-behaviour-aware indices, community risk and societal risk, designed to measure the effect of prioritizing each community for vaccination to reduce (1) its own mortality risks and (2) the mortality risks it imposes on society as a whole. On the basis of these two indices, we design a framework for vaccine prioritization that simultaneously improves social utility and equity in all dimensions across scenarios of different vaccination rates and timing. By providing a general framework to tease out utilitarian and egalitarian values in COVID-19 vaccine distribution, our findings carry broad implications for the design of vaccination strategies.

Results

Behaviour- and demography-informed epidemic modelling. To capture the heterogeneity in health risk faced by different communities^{34–36}, we propose an epidemic model that principally integrates two important factors in the spread of coronavirus—that is, demographic profiles and mobility behaviours (Fig. 1a). Specifically, demographic profile is found to be substantially correlated with the fatality rate of SARS-CoV-2 infection^{28,40}, while mobility behaviour determines the likelihood of exposure to coronavirus^{35,37,39,41}. The proposed behaviour- and demography-informed epidemic model (BD model) therefore divides the studied population on the basis of the minimum geographical units defined by the United States Census Bureau, known as census block groups (CBGs)⁴², and maintains a local SEIR model for each of them to characterize the dynamics of intra-CBG epidemic spread, where the infection-fatality rate (IFR) is adjusted on the basis of the demographic profile and age-specified risks estimated in previous medical research²⁸ (Supplementary Table 1). To capture inter-CBG transmission resulting from urban mobility, the proposed BD model constructs a bipartite network linking CBGs and points of interest (POIs) with time-varying edges to track hourly movements extracted from the SafeGraph dataset⁴³, where the edge weights reflect temporal mobility intensity. New infections occur in POIs and CBGs with different probabilities determined by environmental characteristics, and infected populations travel to other communities proportional to their extracted movements (Methods, ‘Epidemic model, calibration and preliminary analysis’).

We evaluate the proposed BD model in nine large metro statistical areas (MSAs) in the United States covering over 75 million people and compare it with two baseline models: a standard SEIR model and a metapopulation model that only considers heterogeneous mobility among communities³⁵. Results show that the proposed BD model consistently produces more accurate estimations of daily deaths in each MSA with growth patterns ranging from sub-linear to exponential (Fig. 1b). Specifically, the BD model outperforms the SEIR and metapopulation models by reducing 51.9% (95% confidence interval (CI), 0.504–0.534; $P < 0.001$) and 35.7% (95% CI, 0.332–0.382; $P < 0.001$) of the normalized root mean square error (NRMSE), respectively (Fig. 1c and Supplementary Fig. 1). The proposed BD model also reveals higher mortality risks faced by communities with higher older adult ratios, lower household income, higher essential worker ratios and higher minority ratios, consistent with real-world observations^{4,28–30,35,36,39,44} (Fig. 1d). In contrast, the

metapopulation model predicts that communities with higher older adult ratios will face unreasonably lower mortality risk, probably due to its inability to model demographic profiles. Furthermore, the SEIR model cannot capture heterogeneous risks in different communities due to its assumption of homogeneous population mixing.

We examine the correlations between demographic profiles and mobility behaviour to explain heterogeneity across communities predicted by the metapopulation model and our proposed BD model (Fig. 1e). We find that the older adult ratio negatively correlates with per capita mobility ($r = -0.29$), indicating that neglect of demography-specific mortality risk will lead to inaccurate estimations of risk for different age groups. By contrast, our proposed BD model finds that differences in mobility behaviours are outweighed by the change in IFRs due to age structure and predicts higher mortality risk in communities with higher older adult ratios, consistent with previous research^{28,44}. Moreover, communities with lower average household incomes, higher essential worker ratios and higher minority ratios are associated with higher levels of mobility, probably following from limitations in their ability to substantially reduce mobility during the pandemic^{30,35,36}. Both the metapopulation model and the proposed BD model therefore reproduce higher risk associated with low-income communities, while the risk associated with essential worker ratios remains complicated due to the joint effect of demographic and mobility profiles (for example, essential workers generally have higher mobility but younger demographic profiles). In view of this, considering the joint effect of both mobility behaviours and demographic profiles should enable improved prediction of the heterogeneous risks facing different communities. By incorporating both into the epidemic model and utilizing large-scale real-world mobility data for calibration, the proposed BD model is effective in generating accurate daily predictions and capturing heterogeneous risks faced by distinctive communities and provides a framework from which we can analyse the impacts of differing vaccine distribution strategies on social utility and equity.

Consequences of alternative vaccine distribution strategies.

Social utility and equity represent the two most important concerns considered by public health policy makers in the COVID-19 pandemic^{9,14–16}. These account for the collective welfare of the entire society and disparities among individual communities. As a critical policy concern during the COVID-19 pandemic, discussions of vaccine distribution strategies have centred on the trade-off between social utility and equity¹⁵, but this has been inadequately evaluated or supported with empirical data. With the proposed BD model, we aim to reveal mechanisms behind social utility and equity in vaccine distribution with large-scale empirical data. In vaccine distribution, we quantify social utility as a reduction in the overall fatality rate, and we quantify equity as a reduction in the Gini coefficient of fatality rates among communities (Methods, ‘Quantification of social utility and equity’). On the basis of previous analyses that reveal heterogeneous health risks faced by populations with different demographic profiles (Fig. 1d), we focus on four dimensions of equity among communities: equity among age groups, income groups, occupational groups and racial/ethnic groups.

Prioritizing the least advantaged populations is acknowledged as a fundamental value in health-care resource allocation^{46–48}. We therefore construct four vaccine distribution strategies that prioritize the most disadvantaged communities defined on four dimensions of demographic profiles: older adult ratio (Prioritize by Age), average household income (Prioritize by Income), essential worker ratio (Prioritize by Occupation) and minority ratio (Prioritize by Race/Ethnicity). As a baseline for comparison, we also construct a Homogeneous strategy, which provides vaccine access to each community with uniform probability⁴⁹. Our experiments show that strategies that prioritize the worst-off communities drastically improve the equity defined on the corresponding demographic

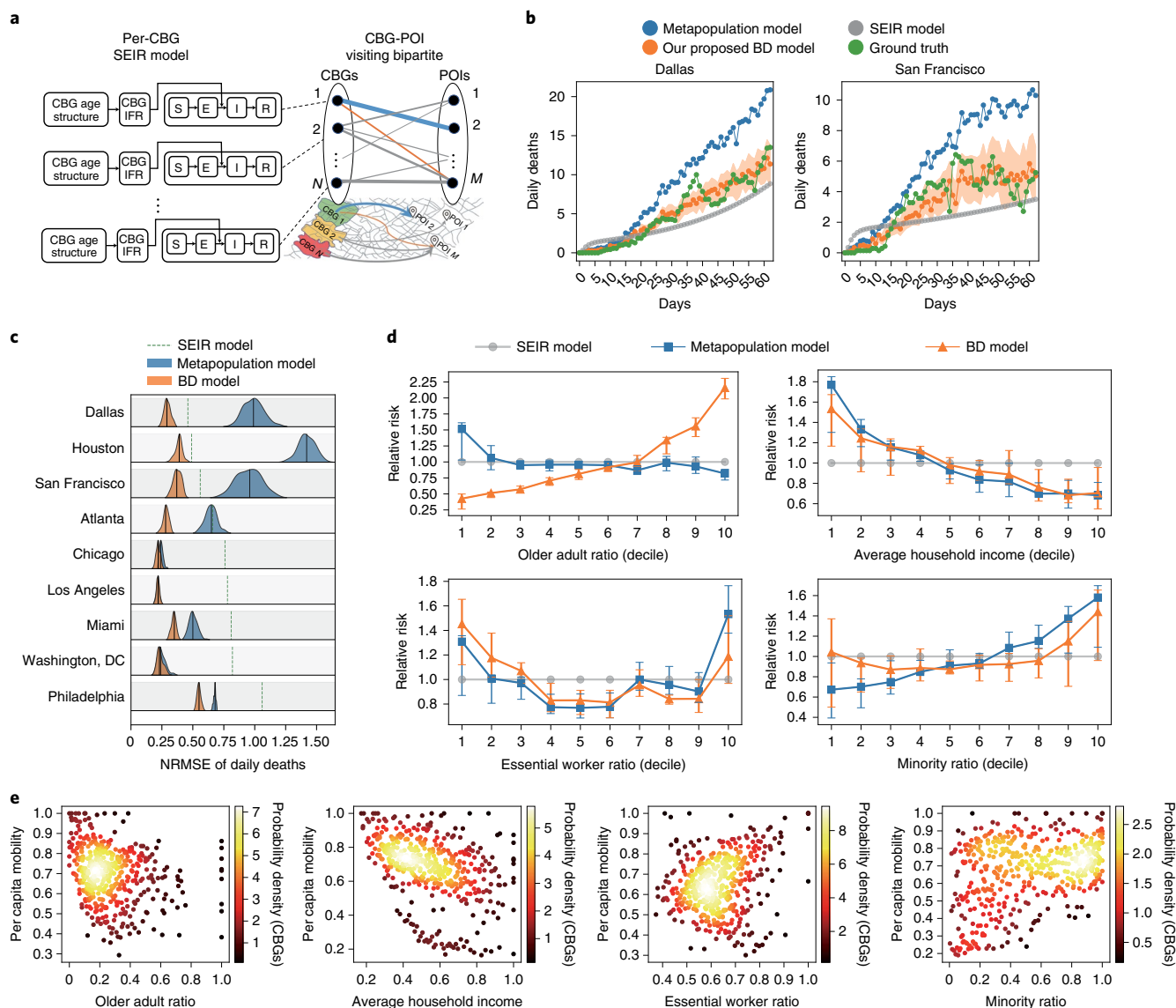


Fig. 1 | Behaviour- and demography-informed epidemic modelling (BD model). **a**, Overview of our BD model, where each CBG maintains its specific SEIR model and connects with other CBGs via mobility flows. **b**, Capability in fitting representative curves of daily deaths. Whether daily deaths grow sub-linearly or almost linearly, our BD model (orange) fits more accurately to ground truth (green) than the SEIR (grey) and metapopulation (blue) models do. The shaded regions show the results of parameter sets that achieve an RMSE within 150% of the best result. **c**, Distribution of NRMSE in daily death prediction (across 100 bootstrap samples). The width of the violin indicates the probability density, and the line within the violin indicates the median value. Our BD model (orange) reduces the error by 51.9% and 35.7% compared with the SEIR model (dashed lines) and the metapopulation model (blue), respectively. **d**, Predicted uneven fatality rates among communities with different demographic features (across 30 independent experiments). The error bars indicate the 25th and 75th percentile values. Our BD model (orange) successfully captures the high mortality risks faced by communities with high older adult ratios, low household income, high essential worker ratios and high minority ratios, while two baseline models fail. **e**, Joint distributions of demographic features and mobility. The older adult ratio and average household income negatively correlate with per capita mobility ($r = -0.29$ and $r = -0.45$, respectively), while the essential worker ratio and minority ratio positively correlate with per capita mobility ($r = 0.39$ and $r = 0.35$).

dimension (Fig. 2a). Specifically, equity defined on age, income, essential worker ratio or minority ratio can be improved by 26.2% (95% CI, 0.168–0.356; $P < 0.001$), 40.5% (95% CI, 0.269–0.541; $P < 0.001$), 43.1% (95% CI, 0.315–0.547; $P < 0.001$) or 31.8% (95% CI, 0.179–0.456; $P < 0.001$) compared with the Homogeneous baseline, respectively. Moreover, in most cases, all four strategies also achieve improvement in social utility compared with the baseline (Fig. 2a), suggesting that overall social utility and equity defined on a specific demographic dimension are likely to be simultaneously improved by prioritizing the most disadvantaged communities within that dimension. This sharply contradicts the

conventional view of an inevitable trade-off between social utility and equity^{19–22}. Detailed analysis reveals that prioritizing communities under greater risk consistently results in larger improvements for both social utility and equity, which further highlights the effectiveness achieved by prioritizing the worst-off communities (Supplementary Table 2).

We further consider and incorporate the potential impacts of vaccine hesitancy and administration capability into our experiments, both of which can substantially undermine the benefit of distributed vaccines. Vaccine hesitancy refers to the phenomenon wherein people mistrust and refuse to take vaccines despite availability,

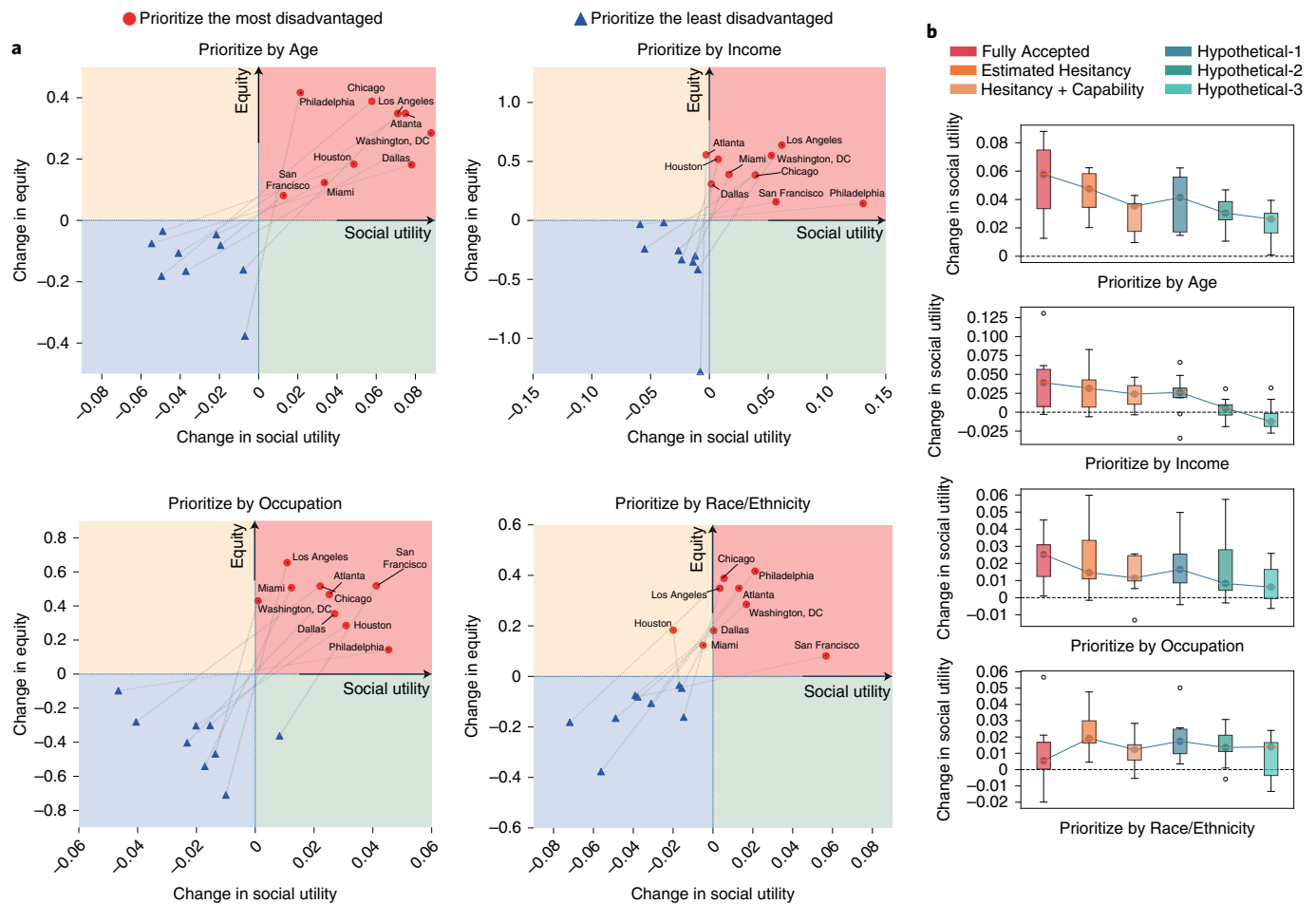


Fig. 2 | Social utility and equity under different vaccine distribution strategies. **a**, Changes in social utility and equity, compared with the Homogeneous baseline. The red and blue points represent the strategies prioritizing the most and least disadvantaged communities, respectively. In each plot, the first to the fourth quadrants represent (1) simultaneously improving utility and equity, (2) improving equity but damaging utility, (3) simultaneously damaging utility and equity, and (4) improving utility but damaging equity. **b**, Change in social utility under different scenarios of vaccine hesitancy (nine MSAs). The bottom and top of each box indicate the 25th and 75th percentile values. The whiskers indicate 1.5 \times the interquartile range below and above the 25th and 75th percentile values. The line inside the box represents the median value of the results. When vaccine hesitancy in low-income communities is stronger, the benefit to social utility brought by prioritizing disadvantaged communities diminishes and is eventually erased, making it inferior to the baseline.

especially in low-income communities^{50,51}. Administration capability refers to the limited capability of local facilities to roll out vaccines, which can be constrained by storage capacity, human force and other institutional factors^{52–54}. First, we construct an Estimated Hesitancy scenario by setting vaccine hesitancy rates for different communities according to a population sample-based national assessment⁵⁰. Second, we construct a Hesitancy+Capability scenario by further superimposing administration capability onto communities, which is estimated from empirical vaccination data provided by the US Centers for Disease Control and Prevention (CDC)⁵⁵. Specifically, administration capability is estimated as the difference between vaccine acceptance hesitancy derived from surveys and the empirical vaccination rate. It captures the percentage of residents willing to take vaccines but who have not done so, probably due to limited administration capacity (for example, no nearby vaccine administration facility). To investigate how different patterns of vaccine hesitancy affect vaccination strategy results, we further design three hypothetical scenarios, where we set vaccine hesitancy rates from the bottom to top income groups at 0.4, 0.3, 0.2, 0.1 and 0 (Hypothetical-1); 0.8, 0.6, 0.4, 0.2 and 0 (Hypothetical-2); and 0.9, 0.7, 0.5, 0.3 and 0 (Hypothetical-3) (Methods, ‘Vaccination scenarios’). These scenarios sequentially reflect larger differences across

income groups and thus allow us to explore more extreme disparities in vaccine hesitancy than we currently observe. In general, improvements in social utility diminish as differences in vaccine acceptance rates grow larger (Fig. 2b). Among the tested prioritization strategies, Prioritize by Income is most sensitive to changes in vaccine hesitancy rates. Nevertheless, its improvement to social utility does not vanish until the vaccine hesitancy rate in the bottom income group rises to five times that in the top group, as in Hypothetical-2, explained by vastly disproportionate risks facing different income groups in the pandemic. This hypothetical hesitancy is far larger than what we observe from data. As expected, in all five scenarios, prioritizing the most disadvantaged communities consistently and significantly improves equity. By following the robust guideline to prioritize the most disadvantaged, social utility and equity can both be improved even if the most disadvantaged groups manifest the most vaccine hesitancy. This demonstrates the out-sized protective impact that would accrue to society from far greater investments in vaccination outreach, education and incentive for our most disadvantaged communities.

For each strategy, we also calculate its impact on equity defined along other demographic dimensions. We find that it is difficult to achieve a comprehensive improvement in all dimensions by

Table 1 | Changes in four dimensions of equity quantified by the Gini index, compared with the Homogeneous baseline

| Policy | Dimension of equity | MSA | | | | | | | | |
|------------------------------|---------------------|-----------|----------|---------|-----------|-------------|----------|--------------|---------------|----------------|
| | | Atlanta | Chicago | Dallas | Houston | Los Angeles | Miami | Philadelphia | San Francisco | Washington, DC |
| Prioritize by age | Age | 34.9%** | 38.8%** | 18.1%** | 18.3%** | 34.9%** | 12.3%** | 41.7%** | 8.1%** | 28.5%** |
| | Income | 18.0% | 7.2%** | 3.1% | 3.1% | (9.0%)* | (6.9%) | (0.9%)** | 0.4% | 2.8% |
| | Occupation | (10.0%)** | 28.6%** | 1.6% | (7.1%) | 4.4% | 2.8% | 1.6%** | (13.3%) | 31.4%** |
| | Race/ethnicity | 28.6%* | (5.2%)** | 14.9%** | 17.1%* | 10.6% | 3.0%** | (1.7%)** | (6.4%) | (39.6%)** |
| Prioritize by income | Age | 5.0% | 3.5%** | 0.9% | 0.9% | (7.7%)** | 2.4%* | (15.7%)** | (0.7%) | (3.8%)** |
| | Income | 55.6% | 38.5%** | 30.8%** | 51.8%** | 63.9%** | 38.9%** | 14.3%** | 15.8%** | 55.0%** |
| | Occupation | 0.0% | 46.3%** | (2.9%) | (0.1%) | 33.8%** | 17.5% | 22.6%** | 6.0% | 22.8%** |
| | Race/ethnicity | (26.7%)** | 29.3%** | (4.4%) | (26.7%)** | (58.4%)** | 7.5%** | 19.2%** | 11.7% | 35.8%** |
| Prioritize by occupation | Age | (5.1%)* | 6.2%** | (4.2%)* | (1.3%) | 0.5% | 1.2% | (12.9%)** | (8.0%)** | 6.9%** |
| | Income | 36.1% | 14.3%** | (6.5%) | (3.6%) | (1.0%) | 4.0% | 4.7%** | 7.4%** | 6.0%* |
| | Occupation | 46.7%** | 51.6%** | 35.4%** | 28.5%** | 65.4%** | 50.7% | 14.3%** | 51.8%* | 43.0%** |
| | Race/ethnicity | (15.6%)* | 12.1%** | (2.9%) | (15.9%) | 23.2%** | 4.5% | 4.9%** | 0.5% | (2.2%) |
| Prioritize by race/ethnicity | Age | 11.3%** | (4.5%)** | 5.3%** | 1.5% | 7.7%** | 4.4%** | (22.3%)** | (5.3%) | (5.8%)** |
| | Income | (29.2%) | 17.8%** | (0.8%) | 0.1% | (14.0%)** | (13.6%)* | 2.8%** | 5.3%* | 23.9%** |
| | Occupation | 0.7% | 14.1%** | (7.1%) | (10.2%) | (6.9%) | 7.4% | 0.4% | 1.0% | (12.4%)** |
| | Race/ethnicity | 45.7%** | 28.8%** | 18.3%** | 27.5%** | 66.8%** | 25.2%** | 8.6%** | 18.6%** | 46.6%** |

Values with both parentheses and asterisks indicate significant degradation in the corresponding dimension of equity. Double-sided paired *t*-test: **P* < 0.05; ***P* < 0.01. For detailed statistics, please refer to Supplementary Table 3.

simply prioritizing the worst-off communities (Table 1). To reveal the relationships between disadvantages along different demographic dimensions, we analyse correlations among the older adult ratio, average household income, essential worker ratio and minority ratio for each CBG (Methods, ‘Correlation analysis of demographic features’). We observe a positive correlation between older adult ratio and average household income ($r=0.14$), indicating that populations with older demographics tend to have higher incomes. Essential worker ratio negatively correlates with older adult ratio ($r=-0.2$) but positively correlates with average household income ($r=0.28$), indicating that populations with larger proportions of essential workers tend to be younger with higher household incomes. Minority ratio negatively correlates with the other three demographic features, but to different extents ($r=-0.31$ with older adult ratio, $r=-0.59$ with average household income and $r=-0.16$ with essential worker ratio; Supplementary Fig. 2). These correlations highlight the mismatch of disadvantaged populations across different demographic dimensions, which results in conflicts between equities that cannot be settled on the basis of demographic features alone. This suggests the need to explore more essential mechanisms underlying demographic features and health that can forecast vaccination outcomes and guide vaccine distribution.

Indices for estimating vaccine prioritization outcomes. To inform the design of vaccine distribution, it is critical to accurately estimate the outcomes that would result from prioritizing certain communities for vaccine distribution. Specifically, for optimal design we must be able to estimate changes in overall social utility and equity when vaccinating each community. Policy designers typically rely either on a single demographic feature⁵⁶ or on indicators computed solely on the basis of demographic data, such as the social

vulnerability index (SVI) designed by the US CDC⁵⁷. Nevertheless, Fig. 2 shows that such approaches will probably degrade equity along certain dimensions, due primarily to complex associations between distinct demographic profiles, infection-fatality risks and per capita mobility (Supplementary Fig. 3). We therefore design two vaccine outcome indices, community risk and societal risk, that capture the underlying mechanisms. To evaluate changes in equity when vaccinating a community, we design a community risk index as the expected mortality rate, calculated as the product of the estimated contact frequency due to average community movement and the infection-fatality risk associated with community demographic profiles. To evaluate changes in risks to society when vaccinating a community, we design a societal risk index as the expected number of deaths caused by infection in that community and the secondary infections they impose on persons from other communities (Methods, ‘Quantification of community risk and societal risk’). Societal risk thus captures the number of lives saved in the whole population by vaccinating certain communities, providing a proxy for social utility. Our proposed community risk and societal risk indices capture different characteristics of communities that could result in trade-offs between social utility and equity. For example, older people with lower mobility face greater mortality risk once infected, but they are less likely to spread the disease than young people with high mobility. This is manifested by their high community risk and low societal risk indices.

We perform regression analysis to examine the power of these indices for estimating outcomes associated with alternative vaccination distribution strategies. Specifically, we generate numerous vaccine distribution instances, each of which vaccinates a set of randomly selected communities covering 2% of the total population. We obtain the impact on social utility and equity

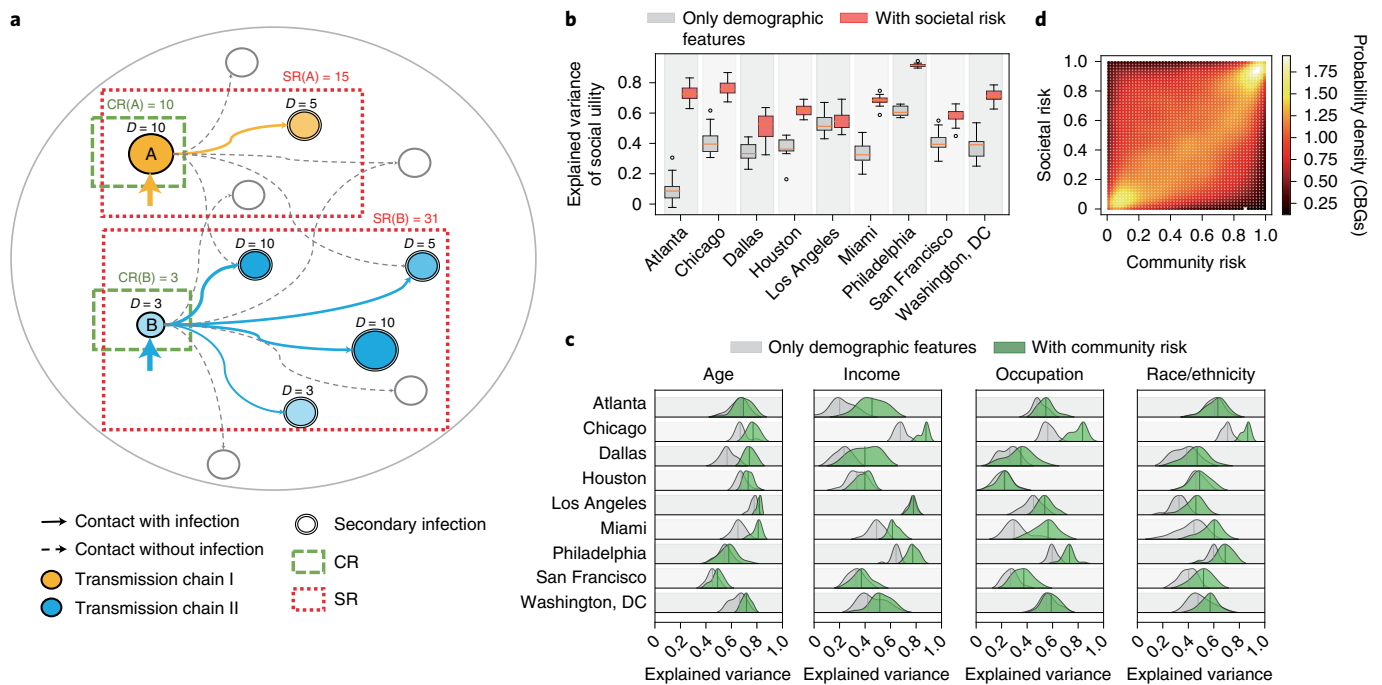


Fig. 3 | Design and justification of community risk and societal risk. a, Illustration of community risk (CR) and societal risk (SR). Each node represents a community, the node size reflects the community's vulnerability and the colour tint reflects the number of deaths in the community, quantified by the value of D . Each edge represents inter-community mobility connections, with thickness reflecting mobility intensity. For each community, CR equals the community's own mortality risk (green boxes), and SR equals the sum of its own mortality risk and the mortality risk it potentially presents to others (red boxes). As two representative cases, community A of transmission chain I has large CR but small SR, while community B of transmission chain II has small CR but large SR. **b**, OLS regression of changes in social utility with and without societal risk (across 20 bootstrap samples). The bottom and top of each box indicate the 25th and 75th percentile values. The whiskers indicate 1.5 \times the interquartile range below and above the 25th and 75th percentile values. Regressions with only demographic features explain on average 38.7% of the variance, measured by adjusted R^2 (grey boxes). The incorporation of societal risk raises the explained variance to an average of 67.5% (red boxes), greatly improving the goodness of fit of the regression model. **c**, OLS regression of changes in equity with and without community risk (across 20 bootstrap samples). The width of the violin indicates the probability density, and the line within the violin indicates the median value. Regressions with only demographic features explain on average 62.9%, 46.0%, 41.4% and 48.5% of the variance, respectively (grey shapes). The incorporation of community risk raises the explained variance to an average of 70.4%, 57.7%, 52.1% and 57.9%, respectively (green shapes), greatly improving the goodness of fit of the regression model. **d**, Joint probability distribution of community risk and societal risk, where brighter colours indicate larger probability density. There is a non-negligible positive correlation ($r=0.29$) between community risk and societal risk.

for each vaccine distribution through simulation, and we perform ordinary least-squares (OLS) regression to estimate changes with demographic features and the proposed indices (Methods, 'Quantification of community risk and societal risk'). Results show that demographic features alone explain only 38.7% (95% CI, 0.365–0.409; $P < 0.001$) of the variance in fatality rate reduction on average, but the incorporation of our societal risk index raises that value to 67.5% (95% CI, 0.656–0.693; $P < 0.001$) (Fig. 3b). In addition, demographic features alone explain on average only 62.9% (95% CI, 0.614–0.643; $P < 0.001$), 46.0% (95% CI, 0.431–0.489; $P < 0.001$), 41.4% (95% CI, 0.391–0.437; $P < 0.001$) and 48.5% (95% CI, 0.465–0.505; $P < 0.001$) of the variances in equity defined on age, income, essential worker ratio and minority ratio, respectively, but incorporating the community risk index raises those values to 70.4% (95% CI, 0.688–0.720; $P < 0.001$), 57.7% (95% CI, 0.549–0.605; $P < 0.001$), 52.1% (95% CI, 0.493–0.548; $P < 0.001$) and 57.9% (95% CI, 0.558–0.599; $P < 0.001$), respectively (Fig. 3c; the detailed regression results are presented in Supplementary Tables 4–12). The indices of societal risk and community risk significantly improve the estimation of changes in social utility and equity under any community prioritization scheme. These two indices also shed light on the simultaneous improvement of social utility and equity with vaccination. Specifically, we discover a positive correlation between community risk and societal risk (Fig. 3d),

which indicates a non-negligible overlap between communities experiencing large community risk and those imposing large societal risk. Therefore, if a vaccine distribution strategy succeeds in targeting such overlapping communities, it can simultaneously achieve improvement in both social utility and equity.

Informing the design of vaccine distribution strategies. On the basis of the proposed indices, we design a flexible framework to generate well-rounded vaccine distribution strategies (that is, a Comprehensive strategy) that can improve both social utility and equity in all demographic dimensions. Our framework integrates community risk, societal risk and demographic profiles with learned weights to generate a comprehensive index of vaccine priority for each community, then distributes vaccines by community accordingly (Methods, 'Design of vaccine distribution strategies'). Besides the Homogeneous baseline and the four strategies examined in Fig. 2, we construct two additional strategies for comparison. First, an SVI-Informed strategy is designed to prioritize vaccines to communities according to the SVI released by the US CDC⁵⁷, which is recommended for use in vaccine prioritization^{56,58}. Second, to further justify the necessity of community risk and societal risk, we construct a Comprehensive-Ablation strategy that utilizes demographic features without our indices (Methods, 'Design of vaccine distribution strategies').

Results show that the Comprehensive strategy yielded by our framework achieves greater improvements in social utility and all dimensions of equity (Fig. 4a–c). Visualizing the epidemic patterns clearly illustrates that this strategy slows down the increase of deaths, flattening the daily death curve (Supplementary Fig. 4). Compared with the No-Vaccination scenario, distributing vaccines with the Homogeneous strategy reduces death rates by only 6.3% (95% CI, 0.049–0.077; $P < 0.001$) on average, while our Comprehensive strategy reduces death rates by 20.9% (95% CI, 0.167–0.252; $P < 0.001$) on average. As for impacts on equity by age, income, occupation and race/ethnicity, the Homogeneous strategy results in little improvement, if not deterioration (–0.1% (95% CI, –0.018 to 0.015; $P = 0.343$), –2.4% (95% CI, –0.072 to 0.025; $P = 0.004$), –0.1% (95% CI, –0.035 to 0.032; $P = 0.096$), 1.8% (95% CI, –0.007 to 0.042; $P = 0.104$)), while our Comprehensive strategy improves these dimensions of equity by 22.5% (95% CI, 0.159–0.292; $P < 0.001$), 33.8% (95% CI, 0.144–0.532; $P < 0.001$), 48.3% (95% CI, 0.304–0.661; $P < 0.001$) and 39.3% (95% CI, 0.212–0.574; $P < 0.001$), respectively. In contrast, all the other prioritization strategies, whether based on demographic features (older adult ratio, average household income, essential worker ratio or minority ratio) or indicators calculated solely from demographic data (SVI), degrade either social utility or certain dimensions of equity and thus fail to strike an optimal balance. Although the Comprehensive-Ablation strategy is informed by the same demographic features as the Comprehensive one, it is still unable to guarantee improvements in all health outcome measures because it does not incorporate the impact of mobility. Demographic features are therefore inadequate to guide the design of vaccine distribution strategies alone, but our proposed vaccination outcome indices (community risk and societal risk) complete the framework and generate strategies that resolve the conflicts among utility and equity values.

To examine the generalizability of this framework, we further construct two sets of experimental scenarios that reflect different levels of vaccine supply and epidemic intensity. To estimate the overall performance of a vaccination strategy, we take the sum of relative changes in social utility and the four dimensions of equity, which approximates the calculation of an L1-norm for a vector, but taking sign into account. In the first set of scenarios, we vary the vaccination rate from 5% to 56% of the total population, reflecting different vaccine supply levels (Fig. 4d). Specifically, the vaccination rate of 56% reflects vaccination progress in the United States by October 2021. In this scenario, we construct an additional strategy, Real-World, which distributes vaccines proportional to the real-world distribution estimated by the US CDC⁵⁵. In the second set of scenarios, we change the timing of vaccination by up to ten days to reflect different levels of epidemic spreading (Fig. 4e). Experiments of finer-grained changes in vaccination rate and timing can be found in the supplementary materials (Supplementary Figs. 5–19). In all variations, our framework successfully yields comprehensive strategies that simultaneously elevate social utility and the four dimensions of equity. We note that the Real-World strategy is highly limited in overall performance, suggesting that there remains substantial space to improve real-world vaccine distribution strategies even with high vaccination rates. Projected improvement is more prominent with lower vaccination rates, which is reasonable because with increased vaccine supply, overlapping vaccinated populations will expand and eventually eliminate the differences between prioritization strategies. Nevertheless, this trend highlights that in the face of greater vaccine shortages, more attention should be paid to coordinate the elevation of overall welfare and the mitigation of health disparities. In sum, our experiments demonstrate that our vaccine distribution framework is generalizable across different MSAs, epidemic burdens and vaccination timings.

Discussion

Coordination among multiple ethical values is of central concern when limited, critical health-care resources such as vaccines must be apportioned to people in the face of profound health crises. In this paper, we aim to strike an improved balance between two critical ethical values, social utility and equity. We note that there are multiple ways to conceptualize inequity in public health undergirded by different theories of distributive justice. To our knowledge, there are at least three distributive justice theories distinguished by their views on what subset of health inequalities should be considered inequity. Cause-oriented theory classifies health inequalities into those caused by nature versus those caused by society and regards only the latter as health inequity⁵⁹. Action-oriented theory shifts focus to avoidability, proposing that health inequity includes health inequalities “amenable to positive human intervention”, regardless of cause¹². This view emphasizes agency-people’s ability to shape the future—acknowledging that humans are able to reduce inequality in many dimensions, even if some result from nature or luck^{10,13,60}. The absolute theory measures health inequity in ungrouped individuals to avoid imposing prejudice and regards virtually all health inequalities as inequity^{20,61}. While each of these theoretical approaches relies on a distinctive mixture of ideological foundations, methodological limitations and pragmatic insights, we build our equity framework primarily on the basis of action-oriented theory. We define health inequity as those health inequalities that could have been mitigated by vaccination, regardless of whether the cause is social or pre-existing health. This research design echoes the framework proposed by National Academies of Sciences, Engineering, and Medicine¹⁵, and it identifies age, income, occupation and race/ethnicity as four critical dimensions of equity in our study.

Different from the traditional view of unavoidable trade-offs between social utility and equity, our BD model reveals that prioritizing the most disadvantaged communities for COVID-19 vaccine access can simultaneously improve social utility and equity. This outcome is driven by underlying community heterogeneity in both mobility behaviour and demographic profile. We resolve the tension of equity across different demographic features by designing two indices, community risk and societal risk, to estimate vaccination outcomes. The effectiveness of both indices reveals the necessity to jointly consider both demographic and behavioural heterogeneity in epidemic modelling and policy design.

The vaccine distribution framework we propose provides clear guidance to policymakers. Currently, the prioritization of vaccines is usually based on a rigid stratification of age or occupation^{38,56,62} and set for the state or country as a whole. In contrast, our framework enables the design of flexible distribution strategies aware of joint effects from mobility behaviours and demographic profiles, which can be tailored to local conditions. Moreover, our framework possesses the following two benefits. First, our method provides meso-scale policy guidance by achieving a balance between effectiveness and ease of implementation. With awareness of heterogeneous risks faced by different CBGs, we can maximize the benefits to society with limited vaccination dosages. Meanwhile, because vaccine priorities are determined on the CBG level, people within the same CBG are not discriminated against, providing local administrations with greater flexibility in the actual vaccine roll-out. Second, our method is privacy-preserving. For both demographic features and mobility records, we use only aggregate data on the CBG level, without revealing any individual information, presenting a minimal invasion of personal privacy. Our framework is therefore not only theoretically informative but also instructive for real-world practice. More broadly, equitable access to immunization is viewed by many as a critical part of the right to health, a fundamental human right endowing every person with the ability to pursue and claim their highest attainable health status^{63,64}. Although it has been listed among the six principles of the Global Vaccine

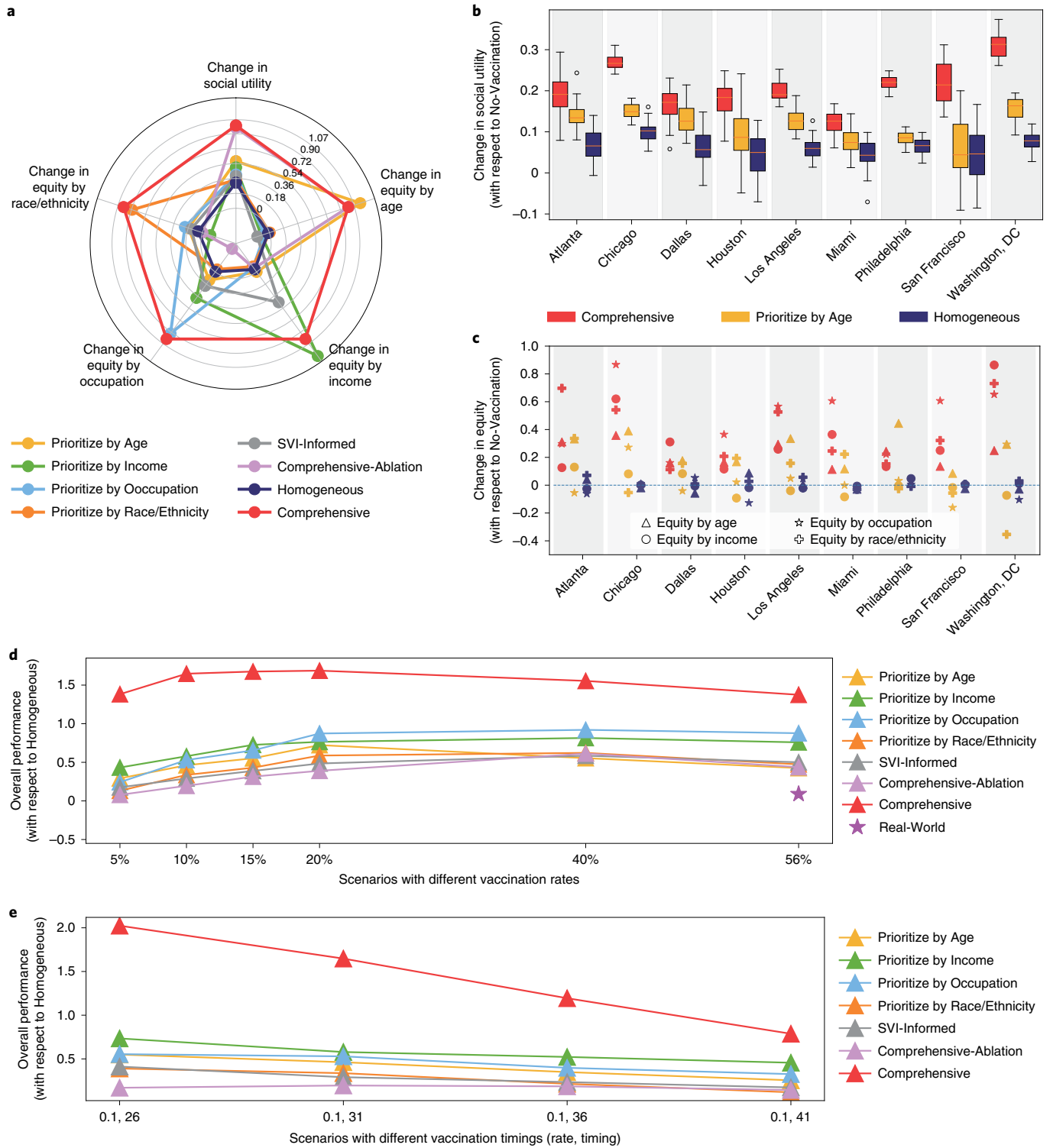


Fig. 4 | Performance of the Comprehensive distribution strategy under various vaccination rates and timings. a, Changes in social utility and four dimensions of equity under eight vaccine distribution strategies. Values are normalized by the result of the Comprehensive strategy. The Comprehensive strategy (red) surpasses or is comparable to all other strategies in the five metrics, indicating its well-rounded effectiveness. In contrast, SVI-Informed (grey) and Comprehensive-Ablation (violet) result in degradation in certain dimensions of equity. **b**, Changes in social utility in each MSA. The bottom and top of each box indicate the 25th and 75th percentile values. Whiskers indicate 1.5x the interquartile range below and above the 25th and 75th percentile values. **c**, Changes in equity by age, income, occupation and race/ethnicity in each MSA. **d**, Overall performance of strategies under different vaccination rates. Overall performance is the sum of relative improvements in social utility and the four dimensions of equity compared with the Homogeneous baseline. The star shows overall performance if a vaccine is distributed proportionally to its real-world distribution, with a vaccination rate of 56% (that is, close to the current rate in the United States). **e**, Overall performance of strategies under different vaccination timings.

Action Plan since 2011⁶⁵, there is still a long way to go before we reach this goal. Our research provides insights to settle concerns regarding ethical values in the global health crisis with experiments on large, real-world data. Our study has several limitations. First, because medical studies on the etiology and pathogenesis of coronavirus are still ongoing, we only consider the widely acknowledged heterogeneity in IFR associated with age. Second, we focus solely on vaccine distribution within a country, but as the COVID-19 pandemic is a global public health emergency, it is of equal necessity to quantitatively study how to coordinate social utility and equity at the international level^{16,48}. Third, our study focuses solely on the demand side without investigating supply-side issues⁶⁶. Future research should focus on the impact of vaccination on people's daily movement and economic recovery; interactions between vaccine manufacturing, transportation and distribution; and vaccine utility and equity issues on a global scale. Nevertheless, our study provides powerful guidance for vaccine prioritization even under circumstances of extreme vaccine hesitancy, recommending far greater societal investments in vaccination outreach, education and incentives for disadvantaged and undervaccinated communities than have hitherto been explored. Vaccinating those worst off represents the best step towards societal protection.

Methods

Epidemic model, calibration and preliminary analysis. Our BD model principally combines demographic profiles and mobility behaviours to simulate epidemic spread in urban communities. First, to reflect the heterogeneity of demographic profiles across communities, we calculate CBG-specific IFRs according to their demographic structure. Here we specifically focus on age structure, as the quantitative influence of age on IFR is the most widely recognized. We divide the population of each CBG into 17 age groups (0–4, 5–9, 10–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79 and 80+), associated with different levels of infection–fatality risk. To evaluate vaccine distribution strategies, we set IFRs according to recent, well-established findings⁵⁸. Because the infection–fatality transition for different individuals is independent, CBG-specific IFRs can be calculated as the weighted average of individual IFRs, where weights are determined by the proportion of each age group in the CBG.

Second, to reflect behavioural heterogeneity, populations from different communities do not homogeneously mix but connect via bipartite mobility networks representing visits from CBGs to POIs each hour, estimated from SafeGraph's data in previous research³⁵. Specifically, each hourly mobility network is represented by a two-dimensional matrix, whose columns correspond to CBGs and rows to POIs. In the simulation, $S \rightarrow E$ transitions take place in both POIs and CBGs, while $E \rightarrow I$ and $I \rightarrow R$ transitions occur only in CBGs. To account for the impacts of different physical environments on transmission characteristics, we use the following adjustable parameters. β_{home} represents the transmission rate within CBGs, which is shared among all CBGs located in the same MSA. β_{poi} represents the basic transmission rate within POIs in the same MSA. For each individual POI, the specific transmission rate $\beta_{\text{poi}(i)}$ is further customized by multiplying β_{poi} by the median dwelling time of visitors (T_i) and the reciprocal of the squared area (S_i^2):

$$\beta_{\text{poi}(i)} = \frac{\beta_{\text{poi}} \times T_i}{S_i^2}. \quad (1)$$

Our simulation period lasts 63 days, which provides us with enough time to simulate the impact of vaccination when the pandemic is at the stage of community transmission¹⁶. Mobility reduction is observed during the simulation period, which reflects the effect of both non-pharmaceutical interventions (such as shop closing) and citizen reactions (such as a reduction in loitering). As our main objective is to tease out the direct outcomes of vaccine distribution strategies, we have not made assumptions about how vaccination in turn impacts mobility.

To introduce flexibility in adapting to different MSAs, we use a vector of age-stratified IFRs to indicate relative risks, and we estimate the actual IFR by fitting a scaling factor specific to each MSA. Such a practice is justified by the consistency in how IFRs increase with age⁴⁰. To calibrate the parameters, we use cumulative fatality records released by the *New York Times* based on reports from state and local health agencies⁶⁷. We first transform cumulative deaths into daily deaths and smooth them with a sliding window of seven days to mitigate randomness. The parameters are optimized by minimizing the RMSE of daily deaths. To mitigate randomness, each experiment is performed in 30 stochastic simulations with different random seeds, the results of which are averaged to obtain the final predictions. To better characterize the distinctive epidemic situations in different MSAs, we do not assume a fixed reproduction number

but instead fit an epidemic model for each. From the well-fitted models, we estimate that the effective reproduction numbers in all MSAs fall around 2, which lies within the plausible range estimated by previous research^{68–70}.

To analyse how mortality risks associate with demographic features, we stratify the whole population into deciles according to their older adult ratio, average household income, essential worker ratio and minority ratio. The fatality rate in each decile is calculated as the total number of deaths divided by the total population in decile. According to the occupation data released by the US Bureau of Labor Statistics, it is estimated that in all states, the lowest essential worker ratio in the labour force is 39.3%⁷¹. Combined with the World Bank's estimate in 2020 that the labour force accounts for about half of the total population⁷², a reasonable essential worker ratio lies beyond 20%. Therefore, when analysing communities stratified by essential worker ratios, we filter those CBGs with a percentage below 20% to mitigate the impact of outliers.

Vaccination scenarios. We assume that vaccines are administered at a single point in time, and we characterize various vaccination scenarios with different vaccination rates (how many vaccines are administered in total) and vaccination timings (when these vaccines are administered). In our first experimental scenario, the vaccination rate is 10%, corresponding to vaccines enough to cover 10% of the population, which is a typical amount of limited supply according to the WHO's SAGE roadmap¹⁶. Vaccination is applied after one month of epidemic simulation (that is, on the 31st day), which reflects the reality in many countries that when vaccines come into use, the epidemic has already reached the stage of community transmission. We assume that all vaccines distributed to communities are effectively used without waste (Fully Accepted), although we alter this in later simulations to model varying levels of vaccine hesitancy.

Corresponding to Fig. 2b, we consider five additional vaccination scenarios to reflect the waste of distributed vaccines due to residents' vaccine hesitancy and communities' limited access to vaccines. In the Estimated Hesitancy and Hesitancy + Capability scenarios, we set vaccine acceptance rates according to a sample-based national assessment⁵⁰. Communities with an average annual household income in the range of [0, 30,000], [30,001, 60,000], [60,001, 99,999] or [99,999, ∞] are associated with a vaccine acceptance rate of 72%, 74%, 81% or 86%, respectively. In Hesitancy + Capability, we further combine our estimated vaccine acceptance rate with real vaccination data from the US CDC⁵⁵, to calculate the vaccine accessibility in each community. The CDC provides the percentage of persons of different ages and ethnic groups that have been fully vaccinated. We take the vaccination data from 15 October 2021 (as shown in Supplementary Tables 13 and 14), calculate the weighted average over CBGs' age structure and ethnicity composition, and multiply the age-determined rates by the ethnicity-determined rates to obtain each CBG's current vaccination rate. If a CBG's vaccination rate is lower than its residents' vaccine acceptance rate, this means that some residents are willing to take vaccines but remain hindered by vaccine accessibility in the CBG. We thus divide obtained vaccination rates by real-world vaccine acceptance rates⁵⁰ to obtain each CBG's vaccine accessibility. In three hypothetical scenarios, we assign different vaccine acceptance rates to five income groups of CBGs to reflect the observed positive correlation between income and vaccine acceptance. Specifically, vaccine acceptance rates from the bottom to top income groups are set to be 0.6, 0.7, 0.8, 0.9 and 1 (Hypothetical-1); 0.2, 0.4, 0.6, 0.8 and 1 (Hypothetical-2); and 0.1, 0.3, 0.5, 0.7 and 1 (Hypothetical-3), sequentially reflecting larger differences across income groups.

Corresponding to Fig. 4d,e and Supplementary Figs. 6–19, we construct two series of vaccination scenarios by varying vaccination rates and vaccination timings to examine the generalizability of our Comprehensive strategy. First, we set the vaccination rate to 5%, 15%, 20%, 40% and 56% of the total population (Fig. 4d) and to 3%, 8%, 13% and 18% of the total population (Supplementary Figs. 6–8 and 12–15). Second, we set the vaccination timing to the 26th, 36th and 41st days (Fig. 4e) and to the 24th, 29th, 34th and 39th days (Supplementary Figs. 9–11 and 16–19).

In all scenarios, we assume that vaccination is fully effective—that is, people that have been vaccinated will not become infected or die from the disease during the remaining time in the simulation. Because we model vaccine distribution on the CBG level, we obtain each CBG's transmission and fatality risks by multiplying their original values by the vaccination rate—that is, if the vaccination rate in a CBG is 50%, the transmission and induced fatality risks its residents experience will halve.

Correlation analysis of demographic features. With demographic data from SafeGraph and the American Community Survey covering more than 42,000 CBGs for nine large MSAs in the United States, we analyse the pairwise correlations among four demographic features: older adult ratio, average household income, essential worker ratio and minority ratio. To eliminate scale differences across demographic features as well as systematic differences across MSAs, we first transform the absolute values of demographic features into percentile ranks for each MSA to reflect their relative levels compared with other CBGs of the same MSA. Next, we aggregate data from different MSAs and calculate the Spearman correlation between each pair of demographic features. In visualizing association patterns between demographic features (Supplementary Fig. 2), the probability

distribution functions are estimated non-parametrically with Gaussian kernel density estimation.

Quantification of social utility and equity. We define metrics of social utility and equity in accordance with the fundamental ethical principles of maximum benefit and mitigation of health inequities in the National Academies of Sciences, Engineering, and Medicine's proposal¹⁵. The maximum benefit principle aims at reducing the total damage of the pandemic. As loss of life represents severe and irreversible damage⁴⁰, we quantify social utility as an overall reduction in the fatality rate. The mitigation of health inequities principle aims to address higher risks faced by certain disadvantaged demographic groups. Specifically, we consider health inequity as health inequalities that could have been mitigated with vaccines, regardless of whether the cause is pre-existing health conditions or social deprivation^{10–13}. The COVID-19 pandemic has imposed substantially higher mortality risk on older adults, low-income households, essential workers and racial/ethnic minorities^{4,30,38,41,44}. Unlike inevitable mortality due to incurable disease, COVID-19-induced mortality risk does not evade prevention and can be effectively reduced with vaccination. We therefore define equity on four demographic dimensions: age, income, occupation and race/ethnicity. First, equity by age is calculated among CBGs with different rates of older adults, where people over 70 years old are defined as 'older adults' in our case, because they face significantly higher risks of death upon infection. Second, equity by income is calculated among CBGs with different average household incomes. Third, equity by occupation is calculated among CBGs with different essential worker ratios. We calculate essential worker ratios by combining the essential occupation list released by Delaware and Minnesota⁷³ with the CBG-level employment data from SafeGraph, following the practice of relevant studies^{41,74}. Finally, equity by race/ethnicity is calculated among CBGs with different rates of racial/ethnic minority population. Here we follow the widely adopted approach to estimate 'racial/ethnic minority' with the complementary percentage of non-Hispanic white residents^{75,76}. More specifically, we adopt the widespread Gini coefficient^{77–80}, which reflects the relative mean absolute difference between all pairs of objects. We take the negative of the Gini coefficient for fatality rates among demographic groups as the inequity in that dimension, calculated as follows: for each demographic dimension, we rank all CBGs according to the corresponding feature and then divide them into N groups covering populations of virtually identical sizes. We denote the fatality rate of the i th group as f_i and the average fatality rate as \bar{f} . The CBG groups are first placed in ascending order of fatality rate. The Gini coefficient G is then calculated as follows:

$$G = \frac{\sum_{i=1}^N (2i - N - 1) f_i}{N \sum_{i=1}^N f_i}. \quad (2)$$

Quantification of community risk and societal risk. The design of the community risk and societal risk indices integrates demography with empirical mobility data. First, the infection risk for each CBG (denoted Φ) is calculated as the sum of two parts: the within-CBG infection risk (Φ_{home}) and the infection risk from POIs that the CBG's residents visit (Φ_{poi}). To obtain the average mobility level for each CBG, we perform calculations on the average CBG–POI visiting matrix over all hours. For each CBG, the average population staying at home is denoted as N_{home} , while the average population visiting the i -th POI is denoted as $N_{\text{poi}(i)}$. Because the effect of age on infection risk is not clearly understood⁸¹, we do not make assumptions about its relation with residents' demography. For residents from any CBG, we denote the within-CBG transmission rate as β_{home} and POI-specific transmission rates as $\beta_{\text{poi}(i)}$.

People staying in CBGs rather than visiting POIs are more likely to stay at home instead of interacting with all others present in the same CBG through uniform mixing. We therefore divide CBG populations by total households according to census data from the American Community Survey to estimate the average household size for each CBG (denoted as $P_{\text{household}}$), and we assume that each individual staying in the CBG makes contact with that average. To deal with outliers, the quotient is further clipped to be at most 10, in accordance with the distribution of household sizes in the United States⁸². The calculations are shown below:

$$\Phi = \Phi_{\text{home}} + \Phi_{\text{poi}} \quad (3)$$

$$\Phi_{\text{home}} = N_{\text{home}} \times P_{\text{household}} \times \beta_{\text{home}} \quad (4)$$

$$\Phi_{\text{poi}} = \sum_i N_{\text{poi}(i)} \times \beta_{\text{poi}(i)}. \quad (5)$$

The fatality rate specific to each CBG is denoted α , and the average fatality rate for all CBGs belonging to the same MSA is denoted $\bar{\alpha}$. For each CBG, community risk is equal to the expected number of deaths among its residents, and societal risk is equal to the expected deaths caused by this CBG across the whole MSA, which consists of two parts: deaths among its own residents and deaths of residents from

other CBGs. In a population with a minority of infectious people but a majority of healthy, susceptible people, the probabilities of being infected and of infecting others are asymmetric, although they depend on the same mobility process. To account for this asymmetry, we further adjust the two terms with the proportion of infected population (denoted γ) and that of susceptible population (denoted δ) obtained by running our simulation until the examined time—that is, the 31st day (Supplementary Table 15). The calculations are shown below:

$$\text{community risk} = \Phi \times \gamma \times \alpha \quad (6)$$

$$\text{societal risk} = \Phi \times \gamma \times \alpha + \Phi \times (\Phi \times \delta) \times \bar{\alpha}. \quad (7)$$

To justify community risk and societal risk as indices of vaccination outcomes, we perform OLS regressions of the change in social utility and the four dimensions of equity in each MSA. In simulating various vaccination strategies, we randomly sample and vaccinate 2% of the total population to achieve a range of vaccination results and then compute health outcomes. To generate more diverse samples in the sampling phase, we first divide the CBGs into $3^6 = 729$ groups, corresponding to their levels (high/median/low) in the four demographic features and the two indices. After merging groups with too few CBGs, we generate a fixed number of samples for each group. For each experiment, we perform 30 stochastic simulations and take the average as our final result.

To examine whether societal risk improves the prediction of the impact of vaccination on social utility, we regress change in social utility on (1) the average and standard deviation of each of the four demographic features (that is, eight independent variables) and (2) the average and standard deviation of each of the four demographic features plus societal risk (that is, ten independent variables). Next, to examine whether community risk improves the prediction of vaccination impact on equity, we regress changes in equity by age, equity by income, equity by occupation and equity by race/ethnicity, respectively, on (1) the average and standard deviation of the four demographic features and (2) the average and standard deviation of the four demographic features plus community risk. To estimate goodness of fit for different regression models, we compare the values of adjusted R^2 , which reflects the proportion of variance in the dependent variable explained by the independent variables, accounting for the number of independent variables.

Design of vaccine distribution strategies. General method of vaccine distribution.

The central step in vaccine distribution is to generate a priority index for each community (CBG) and sequentially distribute vaccines to CBGs according to those priorities. A CBG will not receive vaccines unless those with higher priority are fully vaccinated, so as to prevent discrimination within the same CBG and to focus on CBG-level distribution free from additional hyper-parameters. Specifically, when constructing vaccine distribution priorities according to a single demographic feature (that is, Prioritize by Age, Prioritize by Income, Prioritize by Occupation or Prioritize by Race/Ethnicity), demographic groups are ranked in descending order of predicted fatality rate, and only CBGs belonging to the group with the largest average fatality rate are considered for vaccine access. Inside each demographic group, vaccines are sequentially distributed to CBGs ranked according to the corresponding demographic feature (for example, age). Additionally, an adaptive scheme is introduced to periodically adjust group priorities after a fixed number of vaccines are distributed (for example, 1% of the population).

SVI-Informed vaccine distribution strategy. The SVI is released and maintained by the US CDC / Agency for Toxic Substances and Disease Registry, which combines multiple socio-economic features to assess community resilience in the face of hazardous events, including epidemics⁸⁷. We use this index to construct an SVI-Informed strategy, which distributes vaccines according to SVI ranking in each MSA. We use the current (2020) release of the SVI 2018 for census tracts, which represents the most up-to-date version of the data. Estimated at the level of census tracts, which are larger geographical units than CBGs, SVI does not distinguish CBGs from the same census tract. Accordingly, our SVI-Informed strategy associates CBGs of the same census tract with the same priority and assigns vaccines to them indistinguishably.

Real-World vaccine distribution strategy. Corresponding to Fig. 4d and the scenario with a vaccination rate of 56%, we construct an additional Real-World strategy that distributes vaccines proportionally to the real-world distribution estimated by the US CDC⁸⁵. Following the same methodology as in Methods, 'Vaccination scenarios', we first use vaccination data retrieved in October 2021 to calculate age-and-ethnicity-determined vaccination rates for each CBG. Then, to ensure a fair comparison, we proportionally scale up vaccination rates in all CBGs so that the total number of vaccines is equal to that in other distribution strategies.

Framework for Comprehensive vaccine distribution strategies. Given the complex relationships not only among demographic features but also between demographic features and underlying mechanisms that determine epidemic impact, we must carefully devise strategies for vaccine distribution to obtain simultaneous improvement of social utility and equity along different demographic dimensions.

We therefore propose a flexible framework to automate the design of such a Comprehensive strategy via joint consideration of community risk, societal risk and multiple demographic features. To construct our Comprehensive strategy, we first rank the CBGs according to each of the four demographic features (older adult ratio, average household income, essential worker ratio and minority ratio) and the two indices (community risk and societal risk). We then use TOPSIS, a widely used multi-criteria ranking method³⁸, to obtain a comprehensive index of vaccine priority via a weighted combination of the above six features. The initial weights of the six features are set to be equal. To adapt to the specific demographic and mobility patterns in different MSAs, optimal weights are determined through a greedy process to combine multiple features. Starting from an equally weighted combination, we perform simulations to estimate outcomes in social utility and the four dimensions of equity, and we then adjust the weights according to the following heuristic guidelines: (1) when improvement in any dimension of equity is unsatisfactory, the weight of the corresponding demographic feature will be increased; and (2) when improvement in social utility is unsatisfactory, the framework will require trials to increase the weight of either community risk or societal risk. Pseudocode describing the decision process is provided in Algorithm 1 (Supplementary Note 1).

To justify the critical roles played by the proposed behaviour-and-demography-aware indices, we also construct an ablation of our Comprehensive strategy (Comprehensive-Ablation). In this version, the vaccine priorities of the CBGs are calculated only on the basis of a weighted combination of the four demographic features, removing their community risk and societal risk. We examine the generalizability of our Comprehensive strategy under scenarios with different vaccination rates and timings (see Methods, ‘Vaccination scenarios’ for the details).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

COVID-19 daily death data are available from the *New York Times* (<https://github.com/nytimes/covid-19-data>). Mobile phone mobility data and demographic data for CBGs are freely available upon request for academic, non-profit use from SafeGraph (<https://www.safegraph.com/academics>), which is a data company that aggregates anonymized location data from numerous applications to provide insights about physical places, via the SafeGraph Community. Estimated mobility networks are retrieved from <http://covid-mobility.stanford.edu>. Social vulnerability indices for communities are available at the website of the US Agency for Toxic Substances and Disease Registry (https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html). Vaccination data are available at the website of the US CDC (<https://covid.cdc.gov/covid-data-tracker/#vaccination-demographic>). The processed data have been deposited in Zenodo (<https://doi.org/10.5281/zenodo.6608362>).

Code availability

The code used in this research has been deposited in Zenodo (<https://doi.org/10.5281/zenodo.6609393>).

Received: 26 November 2021; Accepted: 12 July 2022;

Published online: 25 August 2022

References

1. WHO Coronavirus (COVID-19) Dashboard (WHO, accessed 22 August 2021); <https://covid19.who.int/>
2. Wang, Z. & Tang, K. Combating COVID-19: health equity matters. *Nat. Med.* **26**, 458 (2020).
3. Gray, D. M., Anyane-Yeboah, A., Balzora, S., Issaka, R. B. & May, F. P. COVID-19 and the other pandemic: populations made vulnerable by systemic inequity. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 520–522 (2020).
4. Van Dorn, A., Cooney, R. E. & Sabin, M. L. COVID-19 exacerbating inequalities in the US. *Lancet* **395**, 1243–1244 (2020).
5. Padma, T. COVID vaccines to reach poorest countries in 2023—despite recent pledges. *Nature* **595**, 342–343 (2021).
6. Maxmen, A. COVID boosters for wealthy nations spark outrage. *Nature* <https://doi.org/10.1038/d41586-021-02109-1> (2021).
7. ECDC releases COVID-19 vaccination rollout strategies for EU/EEA (ECDC, 2020); <https://www.ecdc.europa.eu/en/news-events/ecdc-releases-vaccination-rollout-strategies-euea>
8. *Vaccine Prioritization Dashboard* (Johns Hopkins Disability Health Research Center, accessed 15 September 2021); <https://disabilityhealth.jhu.edu/vaccine-2/>
9. Toner, E. et al. *Interim Framework for COVID-19 Vaccine Allocation and Distribution in the United States* (Johns Hopkins Center for Health Security, 2020); <https://www.centerforhealthsecurity.org/our-work/publications/interim-framework-for-covid-19-vaccine-allocation-and-distribution-in-the-us>
10. Whitehead, M. The concepts and principles of equity and health. *Health Promot. Int.* **6**, 217–228 (1991).
11. WHO et al. *Monitoring Equity in Health: A Policy-Oriented Approach in Low- and Middle-Income Countries* Tech. Rep. (World Health Organization, 1998).
12. Norheim, O. F. & Asada, Y. The ideal of equal health revisited: definitions and measures of inequity in health should be better integrated with theories of distributive justice. *Int. J. Equity Health* **8**, 40 (2009).
13. Ruger, J. P. Health and social justice. *Lancet* **364**, 1075–1080 (2004).
14. McClung, N. et al. The Advisory Committee on Immunization Practices’ ethical principles for allocating initial supplies of COVID-19 vaccine—United States, 2020. *Morb. Mortal. Wkly Rep.* **69**, 1782–1786 (2020).
15. National Academies of Sciences, Engineering, and Medicine *Framework for Equitable Allocation of COVID-19 Vaccine* (National Academies Press, 2020).
16. WHO SAGE Roadmap for Prioritizing the Use of COVID-19 Vaccines in the Context of Limited Supply: An Approach to Inform Planning and Subsequent Recommendations Based upon Epidemiologic Setting and Vaccine Supply Scenarios, 13 November 2020 Tech. Rep. (World Health Organization, 2020).
17. Bentham, J. *An Introduction to the Principles of Morals and Legislation* (1781); McMaster University Archive for the History of Economic Thought <https://socialsciences.mcmaster.ca/econ/ugcm/3ll3/>
18. Sassi, F., Le Grand, J. & Titmuss, R. Equity versus efficiency: a dilemma for the NHS. *BMJ (Int. Edn)* **323**, 762–763 (2001).
19. Ottersen, T., Mbilinyi, D., Mæstad, O. & Norheim, O. F. Distribution matters: equity considerations among health planners in Tanzania. *Health Policy* **85**, 218–227 (2008).
20. *The World Health Report 2000: Health Systems—Improving Performance* (World Health Organization, 2000).
21. Lindholm, L., Rosen, M. & Emmelin, M. An epidemiological approach towards measuring the trade-off between equity and efficiency in health policy. *Health Policy* **35**, 205–216 (1996).
22. Dolan, P. The measurement of individual utility and social welfare. *J. Health Econ.* **17**, 39–52 (1998).
23. Goldstein, J. R., Cassidy, T. & Wachter, K. W. Vaccinating the oldest against COVID-19 saves both the most lives and most years of life. *Proc. Natl Acad. Sci. USA* **118**, e2026322118 (2021).
24. Castro, M. C. & Singer, B. Prioritizing COVID-19 vaccination by age. *Proc. Natl Acad. Sci. USA* **118**, e2103700118 (2021).
25. Dushoff, J., Colijn, C., Earn, D. J. & Bolker, B. M. Transmission dynamics are crucial to COVID-19 vaccination policy. *Proc. Natl Acad. Sci. USA* **118**, e2105878118 (2021).
26. Buckner, J. H., Chowell, G. & Springborn, M. R. Dynamic prioritization of COVID-19 vaccines when social distancing is limited for essential workers. *Proc. Natl Acad. Sci. USA* **118**, e2025786118 (2021).
27. Bubar, K. M. et al. Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science* **371**, 916–921 (2021).
28. O’Driscoll, M. et al. Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature* **590**, 140–145 (2021).
29. Carrión, D. Neighborhood-level disparities and subway utilization during the COVID-19 pandemic in New York City. *Nat. Commun.* **12**, 3692 (2021).
30. Weill, J. A., Stigler, M., Deschenes, O. & Springborn, M. R. Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. *Proc. Natl Acad. Sci. USA* **117**, 19658–19660 (2020).
31. Bailey, N. T. *The Mathematical Theory of Infectious Diseases and Its Applications* 2nd edn (Hafner/MacMillan, 1975).
32. Nunn, C. & Altizer, S. *Infectious Diseases in Primates: Behavior, Ecology and Evolution* (Oxford Univ. Press, 2006).
33. Li, M. Y. & Muldowney, J. S. Global stability for the SEIR model in epidemiology. *Math. Biosci.* **125**, 155–164 (1995).
34. Mwalili, S., Kimathi, M., Ojiambo, V., Gathungu, D. & Mbogo, R. SEIR model for COVID-19 dynamics incorporating the environment and social distancing. *BMC Res. Notes* **13**, 352 (2020).
35. Chang, S. et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589**, 82–87 (2021).
36. Gozzi, N. et al. Estimating the effect of social inequalities on the mitigation of COVID-19 across communities in Santiago de Chile. *Nat. Commun.* **12**, 2429 (2021).
37. Lai, S. et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* **585**, 410–413 (2020).
38. Fitzpatrick, M. C. & Galvani, A. P. Optimizing age-specific vaccination. *Science* **371**, 890–891 (2021).
39. Mena, G. E. et al. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science* **372**, eabg5298 (2021).
40. Goldstein, J. R. & Lee, R. D. Demographic perspectives on the mortality of COVID-19 and other epidemics. *Proc. Natl Acad. Sci. USA* **117**, 22035–22041 (2020).
41. Glaeser, E. L., Gorbach, C. & Redding, S. J. How much does COVID-19 increase with mobility? Evidence from New York and four other US cities. *J. Urban Econ.* **127**, 103292 (2022).
42. Definition of census block group. *Wikipedia* (accessed 23 August 2021) https://en.wikipedia.org/wiki/Census_block_group

43. *SafeGraph Weekly Patterns* (SafeGraph, accessed 17 August 2021); <https://docs.safegraph.com/docs/weekly-patterns>
44. Davies, B. et al. Community factors and excess mortality in first wave of the COVID-19 pandemic in England. *Nat. Commun.* **12**, 3755 (2021).
45. Jean-Jacques, M. & Bauchner, H. Vaccine distribution—equity left behind? *JAMA* **325**, 829–830 (2021).
46. Persad, G., Wertheimer, A. & Emanuel, E. J. Principles for allocation of scarce medical interventions. *Lancet* **373**, 423–431 (2009).
47. Emanuel, E. J. et al. Fair allocation of scarce medical resources in the time of COVID-19. *N. Engl. J. Med.* **382**, 2049–2055 (2020).
48. Emanuel, E. J. et al. An ethical framework for global vaccine allocation. *Science* **369**, 1309–1312 (2020).
49. MacNaughton, G. Untangling equality and non-discrimination to promote the right to health care for all. *Health Hum. Rights* **11**, 47–63 (2009).
50. Khubchandani, J. et al. COVID-19 vaccination hesitancy in the United States: a rapid national assessment. *J. Community Health* **46**, 270–277 (2021).
51. Dror, A. A. et al. Vaccine hesitancy: the next challenge in the fight against COVID-19. *Eur. J. Epidemiol.* **35**, 775–779 (2020).
52. Carson, S. L. et al. COVID-19 vaccine decision-making factors in racial and ethnic minority communities in Los Angeles, California. *JAMA Netw. Open* **4**, e2127582 (2021).
53. Bunch, L. A tale of two crises: addressing COVID-19 vaccine hesitancy as promoting racial justice. *HEC Forum* **33**, 143–154 (2021).
54. Hou, Z. et al. Assessing COVID-19 vaccine hesitancy, confidence, and public engagement: a global social listening study. *J. Med. Internet Res.* **23**, e27632 (2021).
55. *COVID Data Tracker, Vaccination Demographics* (CDC, accessed 15 October 2021); <https://covid.cdc.gov/covid-data-tracker/#vaccination-demographic>
56. Schmidt, H. et al. Equitable allocation of COVID-19 vaccines in the United States. *Nat. Med.* **27**, 1298–1307 (2021).
57. *CDC/ATSDR Social Vulnerability Index 2018* (CDC, accessed 15 September 2021); https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html
58. *Social Vulnerability & COVID-19* (Michigan Department of Health and Human Services, accessed 15 September 2021); https://www.michigan.gov/documents/coronavirus/Social_Vulnerability_and_COVID-19-v4_715525_7.pdf
59. Rawls, J. *A Theory of Justice* (Harvard Univ. Press, 2020).
60. *Handbook on Health Inequality Monitoring: With a Special Focus on Low- and Middle-Income Countries* (World Health Organization, 2013).
61. Braveman, P. Health disparities and health equity: concepts and measurement. *Annu. Rev. Public Health* **27**, 167–194 (2006).
62. Matrajt, L., Eaton, J., Leung, T. & Brown, E. R. Vaccine optimization for COVID-19: who to vaccinate first? *Sci. Adv.* **7**, eabf1374 (2021).
63. Backman, G. et al. Health systems and the right to health: an assessment of 194 countries. *Lancet* **372**, 2047–2085 (2008).
64. Pillay, N. Right to health and the Universal Declaration of Human Rights. *Lancet* **372**, 2005–2006 (2008).
65. *Global Vaccine Action Plan, 2011–2020* (WHO, 2013).
66. Golan, M. S., Trump, B. D., Cegan J. C. & Linkov, I. *The Vaccine Supply Chain: A Call for Resilience Analytics to Support COVID-19 Vaccine Production and Distribution* (Springer International, 2021).
67. *Coronavirus (COVID-19) Data in the United States* (New York Times, accessed 6 February 2021); <https://github.com/nytimes/covid-19-data>
68. Wilasang, C. et al. Reduction in effective reproduction number of COVID-19 is higher in countries employing active case detection with prompt isolation. *J. Travel Med.* **27**, taaa095 (2020).
69. Al-Raei, M. The basic reproduction number of the new coronavirus pandemic with mortality for India, the Syrian Arab Republic, the United States, Yemen, China, France, Nigeria and Russia with different rate of cases. *Clin. Epidemiol. Glob. Health* **9**, 147–149 (2021).
70. Rowland, M. A. et al. COVID-19 infection data encode a dynamic reproduction number in response to policy decisions with secondary wave implications. *Sci. Rep.* **11**, 10875 (2021).
71. *US States with the Most Essential Workers* (United Way, accessed 23 August 2021); <https://unitedwaynca.org/stories/us-states-essential-workers/>
72. *Labor Force in the United States* (World Bank, accessed 23 August 2021); <https://data.worldbank.org/indicator/SL.TLF.TOTL.IN?locations=US>
73. *Delaware List of Essential and Non-essential Businesses* (WPVI, 2020); <https://dig.ablocal.com/wpvi/pdf/Delaware-list-of-essential-and-nonessential-businesses-March-22-2020.pdf>
74. Sanchez, D. G., Parra, N. G., Ozden, C. & Rijkers, B. *Which Jobs Are Most Vulnerable to COVID-19? What an Analysis of the European Union Reveals* Res. and Policy Brief No. 148384 (World Bank, 2020).
75. Van Voorhees, B. W., Walters, A. E., Prochaska, M. & Quinn, M. T. Reducing health disparities in depressive disorders outcomes between non-Hispanic whites and ethnic minorities. *Med. Care Res. Rev.* **64**, 157S–194S (2007).
76. Fujishiro, K. & Koessler, F. Comparing self-reported and O*NET-based assessments of job control as predictors of self-rated health for non-Hispanic whites and racial/ethnic minorities. *PLoS ONE* **15**, e0237026 (2020).
77. Illsley, R. & Le Grand, J. *The Measurement of Inequality in Health* (Palgrave Macmillan UK, 1987).
78. Leclerc, A., Lert, F. & Fabien, C. Differential mortality: some comparisons between England and Wales, Finland and France, based on inequality measures. *Int. J. Epidemiol.* **19**, 1001–1010 (1990).
79. Berndt, D. J., Fisher, J. W., Rajendrababu, R. V. & Studnicki, J. Measuring healthcare inequities using the Gini index. In *Proc. 36th Annual Hawaii International Conference on System Sciences* (ed. Dennis, E.) <https://doi.org/10.1109/HICSS.2003.1174353> (IEEE, 2003).
80. Dixon, P. M., Weiner, J., Mitchell-Olds, T. & Woodley, R. Bootstrapping the Gini coefficient of inequality. *Ecology* **68**, 1548–1551 (1987).
81. Dudley, J. P. & Lee, N. T. Disparities in age-specific morbidity and mortality from SARS-CoV-2 in China and the Republic of Korea. *Clin. Infect. Dis.* **71**, 863–865 (2020).
82. *Distribution of Households in the United States from 1970 to 2020, by Household Size* (US Census Bureau, 2020); <https://www.statista.com/statistics/242189/distribution-of-households-in-the-us-by-household-size/>
83. Tzeng, G.-H. & Huang, J.-J. *Multiple Attribute Decision Making: Methods and Applications* (CRC, 2011).

Acknowledgements

This work was supported in part by the National Key Research and Development Program of China under grant no. 2020AAA0106000 to Y.L. and the National Natural Science Foundation of China under grant no. U1936217 to Y.L. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

F.X., P.H., Y.L. and J.E. jointly launched this research and provided the research outline. L.C., F.X. and Y.L. designed the research methods. L.C. performed the experiments and prepared the figures. F.X., Z.H., K.T., Y.L., P.H. and J.E. provided critical revisions. All authors jointly analysed the results and participated in writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41562-022-01429-0>.

Correspondence and requests for materials should be addressed to Fengli Xu, Pan Hui, James Evans or Yong Li.

Peer review information *Nature Human Behaviour* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted <i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> | Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection No software was used for data collection.

Data analysis Data analysis was performed using Python (version 3.6.12). The data is available at <https://doi.org/10.5281/zenodo.6608362>. The source code is available at: <https://doi.org/10.5281/zenodo.6609393>.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

COVID-19 daily death data are available at the New York Times (<https://github.com/nytimes/covid-19-data>). Mobile phone mobility data and demographic data for census block groups are freely available upon request for academic, non-profit use from SafeGraph (<https://www.safegraph.com/academics>), which is a data company that aggregates anonymized location data from numerous applications in order to provide insights about physical places, via the SafeGraph Community. Estimated mobility networks are retrieved from <http://covid-mobility.stanford.edu>. Social vulnerability indices for communities are available at the website of U.S. Agency for Toxic Substances and Disease Registry (<https://www.atsdr.cdc.gov/>)

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|-------------------|--|
| Study description | Based on mobility data and demographic data, we designed an epidemic model, and used it to analyze the outcomes of alternative vaccine distribution strategies on social utility and equity. We propose two indices to capture the impact from vaccinating a certain neighborhood. We propose a framework to design vaccine distribution strategies that can simultaneously improve social utility and equity. |
| Research sample | We study existing data on nine of the largest American metropolitan statistical areas that have consistent death counting method during the period. COVID-19 daily death data are collected from the New York Times (https://github.com/nytimes/covid-19-data). Mobile phone mobility data and demographic data for census block groups are collected from SafeGraph (https://www.safegraph.com/academics). Estimated mobility networks are collected from http://covid-mobility.stanford.edu . Social vulnerability indices for communities are collected from the website of U.S. Agency for Toxic Substances and Disease Registry (https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html). Vaccination data are collected from the website of U.S. CDC (https://covid.cdc.gov/covid-data-tracker/##vaccination-demographic). We did not perform sampling as they were comprehensive. Safegraph mobility data have been shown to be representative of the population (https://docs.safegraph.com/docs/faqs#section-what-is-the-breakdown-of-safe-graphs-mobility-device-panel-is-there-a-skew-toward-certain-demographics). Other data were designed to cover the entire population. The median age and median household income in MSA Atlanta, Chicago, Dallas, Houston, Los Angeles, Miami, Philadelphia, San Francisco, and Washington is 37, 37, 32, 32, 37, 42, 37, 37, 37 and 58820, 62683.5, 60050, 56065, 67321, 52083, 65667, 93571, 96506, respectively. These statistics distribute around the national average of 38.2 and 64324, suggesting the data collection is representative. |
| Sampling strategy | We did not perform sampling, but included all data. |
| Data collection | We did not collect data, but utilized existing datasets. |
| Timing | 2020-03-01 ~ 2020-05-02 |
| Data exclusions | No data were excluded from analysis. |
| Non-participation | No participants were involved in the study. |
| Randomization | This study does not include randomized controlled trials, and data were not organized into experimental groups. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| n/a | Involved in the study |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

| n/a | Involved in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |